

Claims

- [c1] 1. A system prefetching and replacing pages in storage, said storage retaining a plurality of pages, each of said pages comprising a plurality of nodes grouped into one or more regions, said system comprising:
 - (a) a memory management system storing a variable set of pages in memory;
 - (b) a prefetcher recognizing access patterns and usage and prefetching pages among said plurality of pages that fit said access patterns and usage; and
 - (c) a page replacer working in conjunction with said memory management system, and during a traversal, weighting said variable set of pages to identify a subset to be retained and a remainder to be replaced, said subset including pages having a high probability of being revisited and said remainder replaced with a page corresponding to said traversal, and said weighting based upon one or more of the following numerical values associated with each page in said variable set of pages: number of children, number of parents, and region statistics.
- [c2] 2. A system as per claim 1, wherein said weighting is ad-

ditionally based upon identifying pivot pages defining traversals that are not strictly parent-to-child or child-to-parent.

- [c3] 3.A system as per claim 1, wherein said region statistics are any of, or a combination of, the following: minimum step, minimum level, maximum step, or maximum level.
- [c4] 4.A system as per claim 1, wherein said plurality of nodes are associated with a mark-up language based document.
- [c5] 5.A system as per claim 4, wherein said mark-up based language is XML.
- [c6] 6.A system as per claim 1, wherein said regions are node descendant regions.
- [c7] 7.A system as per claim 1, wherein said system is implemented across networks.
- [c8] 8.A system as per claim 7, wherein said across network element is any of the following: local area network (LAN), wide area network (WAN), the Internet, cellular network, or wireless network.
- [c9] 9.A method for robustly prefetching and replacing pages in a system storing a plurality of pages, each of said stored pages comprises a plurality of nodes grouped into

one or more regions, said method comprising the steps of:

(a) storing a variable set of pages in memory;

(b) recognizing access patterns and usage and prefetching pages among said plurality of pages that fit said access patterns and usage;

(c) upon traversals within said plurality of pages:

(i) retaining a subset of said variable set to include pages having a high probability of being revisited; and

(ii) dynamically replacing remainder of said variable set with a page corresponding to said traversal;

wherein, during each of said traversals, said variable set of pages is weighted to identify said subset to be retained and said remainder to be replaced, said weighting based upon at least the following numerical values associated with each page in said variable set of pages: number of children, number of parents, and region statistics.

[c10] 10. A method as per claim 9, wherein said region statistics are any of, or a combination of, the following: minimum step, minimum level, maximum step, or maximum level.

[c11] 11. A method as per claim 9, wherein said weighting is additionally based upon identifying pivot pages defining traversals that are not strictly parent-to-child or child-to-parent.

- [c12] 12.A method as per claim 9, wherein said plurality of nodes are associated with a mark-up language based document.
- [c13] 13.A method as per claim 12, wherein said mark-up based language is XML.
- [c14] 14.A method as per claim 9, wherein said regions are node descendant regions.
- [c15] 15.A method as per claim 9, wherein said method is implemented across networks.
- [c16] 16.A method as per claim 15, wherein said across network element is any of the following: local area network (LAN), wide area network (WAN), the Internet, cellular network, or wireless network.
- [c17] 17.An article of manufacture comprising computer usable medium having computer readable program code embodied therein for robustly prefetching and replacing pages in a transactional system storing a plurality of pages, each of said stored pages comprises a plurality of nodes grouped into one or more regions, said medium comprising:
 - (a) computer readable program code instructing a computer to store a variable set of pages in memory;

(b) computer readable program code recognizing access patterns and usage and fetching pages among said plurality of pages that fit said access patterns and usage;

(c) upon traversals within said plurality of pages:

(i) computer readable program code instructing said computer to retain a subset of said variable set to include pages having a high probability of being revisited; and

(ii) computer readable program code instructing said computer to dynamically replace remainder of said variable set with a page corresponding to said traversal; wherein, during each of said traversals, computer readable program code weights said variable set of pages to identify said subset to be retained and said remainder to be replaced, said weighting based upon at least the following numerical values associated with each page in said variable set of pages: number of children, number of parents, minimum step, and minimum level.

[c18] 18. An article of manufacture as per claim 17, wherein said weighting is additionally based upon identifying pivot pages defining traversals that are not strictly parent-to-child or child-to-parent.

[c19] 19. An article of manufacture as per claim 17, wherein said plurality of nodes are associated with a mark-up language based document.

[c20] 20. An article of manufacture as per claim 19, wherein said mark-up based language is XML.

[c21] 21. A system for prefetching and replacing pages in storage, said storage retaining a plurality of pages, said plurality of pages, at any given point in time, comprising a page pinned in memory, a set of pages unpinned in memory, remainder pages on said disk, each of said pages comprising a plurality of nodes grouped into one or more regions, said system comprising:

- (a) a memory management system comprising said pinned page and said set of unpinned pages;
- (b) a prefetcher recognizing access patterns and usage and prefetching pages among said plurality of pages that fit said access patterns and usage; and
- (c) a page replacer working in conjunction with said memory management system, and during each traversal: updating a pointer to said pinned page to point to a newly pinned page corresponding to said traversal and unpinning previously pinned page, and weighting said previously pinned page and current set of unpinned pages to identify pages to be retained and a page to be replaced, said weighting identifying pages having a high probability of being revisited during future traversals and said weighting based upon at least the following numerical values associated with each page in

said variable set of pages: number of children, number of parents, traversals that are not strictly parent-to-child or child-to-parent, and region statistics.

- [c22] 22.A method as per claim 21, wherein said plurality of nodes are associated with a mark-up language based document.
- [c23] 23.A method as per claim 22, wherein said mark-up based language is XML.
- [c24] 24.A method as per claim 21, wherein said regions are node descendant regions.
- [c25] 25.A method as per claim 21, wherein said method is implemented across networks.
- [c26] 26.A method as per claim 25, wherein said across network element is any of the following: local area network (LAN), wide area network (WAN), the Internet, cellular network, or wireless network.
- [c27] 27.An article of manufacture comprising computer usable medium having computer readable program code embodied therein prefetching and replacing pages in storage, said storage retaining a plurality of pages, said plurality of pages, at any given point in time, comprising a page pinned in memory, a set of pages unpinned in

memory, remainder pages on said disk, each of said pages comprising a plurality of nodes grouped into one or more regions, said system comprising:

- (a) computer readable program code, in conjunction with a computer, implementing a memory management system comprising said pinned page and said set of unpinned pages;
- (b) computer readable program code, in conjunction with said computer, implementing a prefetcher recognizing access patterns and usage and prefetching pages among said plurality of pages that fit said access patterns and usage; and
- (c) computer readable program code, in conjunction with said computer, implementing a page replacer working in conjunction with said memory management system, and during each traversal:

computer readable program mode updating a pointer to said pinned page to point to a newly pinned page corresponding to said traversal ;

computer readable program code unpinning previously pinned page; and

computer readable program code weighting said previously pinned page and current set of unpinned pages to identify pages to be retained and a page to be replaced, said weighting identifying pages having a high probability of being revisited during future traversals and said

weighting based upon at least the following numerical values associated with each page in said variable set of pages: number of children, number of parents, traversals that are not strictly parent-to-child or child-to-parent, and region statistics.

- [c28] 28. An article of manufacture as per claim 27, wherein said plurality of nodes are associated with a mark-up language based document.
- [c29] 29. An article of manufacture as per claim 28, wherein said mark-up based language is XML.